Decoding the First-and-Last Mile: Analyzing the Transit Connecting Bike Share GPS Routes in Hamilton, Ontario with Spatiotemporal Distance Decay

Zehui Yin* (yinz39@mcmaster.ca) Darren M. Scott*

*TransLAB (Transportation Research Lab), School of Earth, Environment & Society, McMaster University

Background and Motivation

- Traditional fixed route public transit systems have the first-and-last mile (FM/LM) problem, as transit stops are never the actual trip origin or destination.
- Shared micromobility services, such as bike share, have the potential to complement public transit by serving as an FM/LM solution.
- Shared micromobility trip data, such as origin and destination pairs, do not contain information about whether a trip is connecting with transit.

Research Questions

- What are the spatiotemporal patterns and trip attributes of First Mile (FM) and Last Mile (LM) bike share trips?
- How do these patterns and attributes differ between FM, LM, and all bike share trips?

Data and Methods

We integrate Hamilton Bike Share GPS tracks (September 3rd to December 30th, 2023), Points of Interest (POI) data from Scholars GeoPortal, Hamilton Street Railway (HSR) GTFS static schedules, and park data from Open Hamilton to build a framework for identifying bike-to-transit connections. GPS tracks are map-matched using a Hidden Markov Model with custom pre- and post-processing in Python, achieving a 99.7% fully match rate. Two travel time matrices, measured in network walking distance (meters) based on OpenStreetMap, were computed using the *r5r* package in R. These matrices link each bike share trip's origin and destination to all HSR transit stops and POIs in Hamilton. Temporal distance is calculated as the time difference (in minutes) between the trip start or end time and the nearest bus arrival at the corresponding transit stop, capturing the potential transfer time of a trip being LM or FM, respectively.

Spatiotemporal distance decay with built environment adjustment formulation (parameter choices informed by existing literature and sensitivity analysis to calibrate the overall number of identified trips to match the 2016 Transportation Tomorrow Survey data):

$$Prob_i = \begin{cases} \left[1 - \left(\frac{d_i}{100}\right)^2\right]^2 \cdot \left[1 - \left(\frac{t_i}{15}\right)^2\right]^2 & \text{when } d_i \leq 100 \text{ and } t_i \leq 15 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$Prob = \max\left\{\underbrace{Prob_1, \cdots, Prob_i, \cdots, Prob_k}_{\text{15 nearest bus arrival}}\right\} \quad (2)$$

$$POI_j = \begin{cases} \left[1 - \left(\frac{d_j}{150}\right)^2\right]^2 & \text{when } d_j \leq 150 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$POI = \sum_{j=1}^{n} POI_j \quad (4)$$

$$Prob_{adj} = \max\left\{Prob \underbrace{-1 \cdot POI - 0.5 \cdot I(parks)}_{\text{penalty}}, \ 0\right\} \quad (5)$$

Analysis Results

For each bike share trip, our framework assigns two values, $\{Prob_{adj,FM}, Prob_{adj,LM}\}$, each ranging from 0 to 1, representing the probability of the trip being a FM or LM connection to transit, respectively. Using these values as weights, we analyze the temporal patterns, trip attributes, and spatial distributions of FM/LM trips as well as all bike share trips. A constant weight was assigned to all trips when examining overall trip patterns.

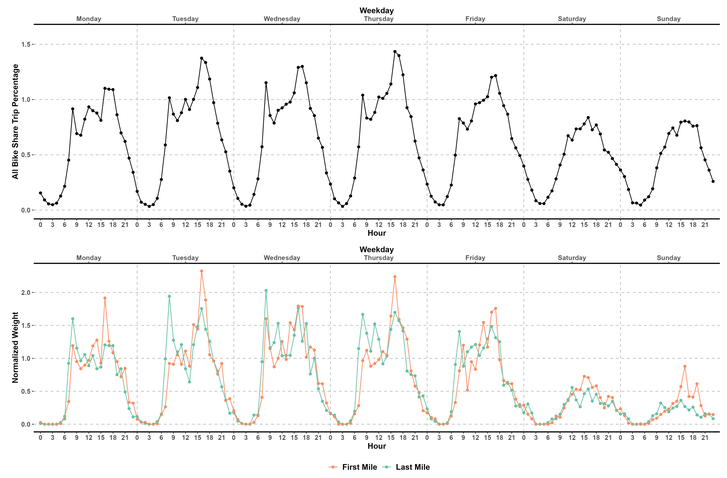

Figure 1. Weekly temporal distribution of trip start times for FM/LM trips (bottom figure) and all bike share trips (top figure) in Hamilton.

Table 1. All and FM/LM Trip Attributes

Attributes	All Bike Share Trips	First Mile Trips	Last Mile Trips
	(Weighted) Mean		
Distance in Meters	1883.85	1590.76	1581.93
Duration in Minutes	11.13	8.29	8.59
Number of Left Turn	4.36	4.03	3.73
Number of Right Turn	4.36	3.93	3.88
Number of U Turn	0.33	0.28	0.22
	(Weighted) Median		
Route Directness Ratio	1.35	1.33	1.33

Table 2. All and FM/LM Trip Attributes Statistical Tests

Attributes	All Bike Share vs. First Mile Trips	All Bike Share vs. Last Mile Trips	First Mile vs. Last Mile Trips
	Weighted Mean Test Bootstrapping p-values		
Distance in Meters	0	0	0.77
Duration in Minutes	0	0	0.39
Number of Left Turn	0	0	0
Number of Right Turn	0	0	0.50
Number of U Turn	0	0	0
	Weighted Median Test Bootstrapping p-values		
Route Directness Ratio	0.01	0	0.81

To compute p-values, samples were centered by aligning weighted means or medians. Then, 1000 bootstrap pairs (matching original sample sizes) were generated using trip weights. Mean/median differences were computed, and two-sided p-values were calculated based on the proportion of bootstrap differences exceeding the observed difference.

We merge overlapping map-matched GPS traces and either count the number of overlaps or sum the FM/LM weights to examine their spatial patterns at the road level.

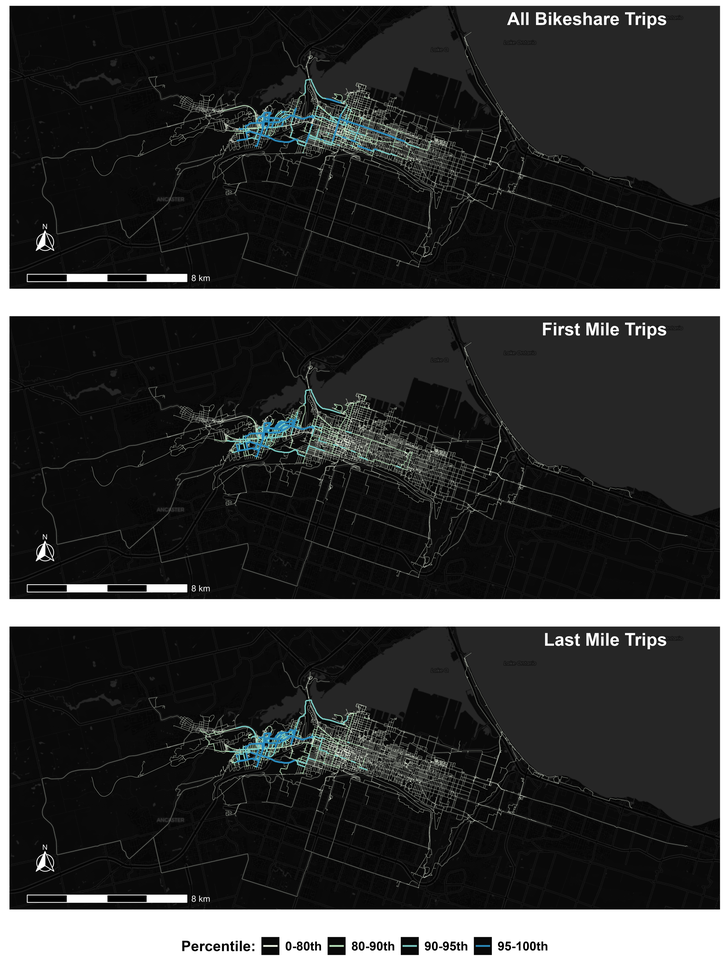

Figure 2. Spatial distribution of FM/LM trips (middle and bottom figures) and all bike share trips (top figure), with routes map-matched in Hamilton.

Main Findings

- FM/LM trips are mostly utilitarian travel, more likely to occur during typical commute peak hours.
- There are more LM trips during the morning peak, while there are more FM trips during the afternoon peak.
- West side of the service area is the most heavily used, with all the most heavily used roads falling within the service area.
- Downtown usage is noticeably lower for both FM and LM trips compared to all bike share trips.
- FM/LM trips are relatively shorter and slightly more direct compared to bike share trips overall.
- FM trips are more challenging compared to LM trips due to a higher number of left and U-turns.

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC)